

Comment on “Sexual Maturation in Relation to Polychlorinated Aromatic Hydrocarbons: Sharpe and Skakkebaek’s Hypothesis Revisited”

The paper of Den Hond et al. (2002) is very interesting, mainly because the reported work demonstrates the potential of biomarkers in environmental health programs. However, we argue with the authors’ conclusion that their findings

suggest that, in line with the concept of endocrine disruption and Sharpe and Skakkebaek’s hypothesis [Sharpe and Skakkebaek 1993], environmental exposure to PCAHs [polychlorinated aromatic hydrocarbons] may adversely interfere with the sexual maturation during the fetal and pubertal stages of development.

The description of the project suggests that the study was designed to investigate the effects of incinerator emissions on the local population in two suburbs of Antwerp, Belgium, Wilrijk and Hoboken. However, the study was part of a larger research project commissioned by the Flemish government to determine the feasibility and issues to be considered in developing environmental health monitoring. More particular, the study of Den Hond et al (2002) was designed to determine whether evaluation of biomarkers measured as part of the regular medical examination of adolescents at the end of their secondary education period would be feasible. We were asked by the Flemish government to review the results of this project (Cuijpers et al. 2000; Ghoetghebeur et al. 2000). We feel that the objectives and hypotheses underlying a feasibility study should not be convoluted with those relative to the hypotheses arising from specific concerns in certain *a priori*-defined regions.

A major shortcoming of the design is the lack of randomness regarding both the study areas and the recruitment of the participants. The study areas were well-known polluted regions within the conurbation of Antwerp. This knowledge may have influenced the choice of tests to be performed. Researchers should be cautious regarding *a priori* choices of study areas, whether or not the sponsor influences the choice.

The use of volunteers instead of randomly selected participants is another flaw, introducing an extra risk of confounding. As a consequence, for example, the proportion of boys to girls in the three study areas differed considerably: from 0.7 to 3.0. In this case there may have been a selection bias of adolescents in lesser physical condition who volunteered for examination.

Reported results were the outcome of analyzing a multitude of associations among

the empirical data; they did not emerge after a rigorous test of an *a priori*-formulated hypothesis, at least not to our knowledge. Such findings are valuable because they may generate hypotheses for further research. In this respect we agree with the authors’ closing sentence (Den Hond et al. 2002):

... [F]urther studies should be undertaken to confirm or to refute our interpretation of the present findings.

The sexual development of the study subjects was judged by the examining physicians. The correspondence in ratings varied between fair and good (as validated by kappa coefficients) (Fleiss 1981). It is then important to exclude a systematic over- or underestimation by one of the examining physicians. This test was not been reported in the study.

Finally, we would like to remark on the reference the authors made (Den Hond 2002) to an earlier report on pregnancy outcome in Flanders (Aelvoet et al 1999):

... [I]n 1997 the Flemish government reported a higher percentage of medically assisted conceptions in the district around the waste incinerators compared with the rest of Flanders....

We feel that this report is incorrectly and selectively referenced, suggesting that the findings are in line with earlier results. Taking into account the mother’s age, no significant differences in way of delivery were found in the earlier study (Aelvoet et al 1999). Furthermore, that report does not contain statements about the necessity of medical assistance at delivery. Part of the above criticism was outlined in the review reports (Cuijpers et al. 2000) but unfortunately not addressed in the article by Den Hond et al. (2002).

To conclude, we would like to reiterate that we commend the authors with their highly relevant and extensive research work. However, they should have placed their study more clearly in the perspective of the Flemish environmental health project. Their conclusions generate interesting hypotheses for further work but should not be viewed as the results of a cause-and-effect study.

Geert Molenberghs

Center for Statistics
Limburgs Universitair Centrum
Diepenbeek, Belgium

Carin Cuypers

Rijksinstituut voor Volksgezondheid en Milieu
Bilthoven, The Netherlands

Els J.T. Goetghebeur

Vakgroep Toegepaste Wiskunde en Informatica
Universiteit Gent
Gent, Belgium

Wim F. Passchier

Department of Health Risk Analysis
and Toxicology
Universiteit Maastricht
Maastricht, The Netherlands
E-mail: wf.passchier@grat.unimaas.nl

John Pieters

Retired officer of the State Inspectorate
of Health
The Hague, The Netherlands

REFERENCES

Aelvoet W, Nelen V, Schoeters G, Vanoverloop J, Wallijn E, Vlietinck R. 1999. Gepercipieerd exces aan congenitale afwijkingen en niet gepercipieerd exces aan meerlingen in de buurt van twee huisvuilverbrandingsovens in Neerland. In: Gezondheidsindicatoren 1997 (Aelvoet W, Fortuin M, Hoofft P, Vanoverloop J, eds). Brussels: Ministerie van de Vlaamse Gemeenschap, p. 130–141.

Cuijpers C, Molenberghs G, Passchier WF, Pieters JLL. 2000. De vinger aan de pols van een gezond Vlaams milieu. (2). Rapportage van een review-commissie van het Project ‘Milieu en Gezondheid’ [in Dutch]. Brussels:Ministerie van de Vlaamse Leefgemeenschap.

Den Hond E, Roels HA, Hoppenbrouwers K, Nawrot T, Thijs L, Vandermeulen C, et al. 2002. Sexual maturation in relation to polychlorinated aromatic hydrocarbons: Sharpe and Skakkebaek’s hypothesis revisited. Environ Health Perspect 110:771–776.

Fleiss JL. 1981. Statistical Methods for Rates and Proportions. New York:John Wiley & Sons.

Ghoetghebeur EJT, Passchier WF, Pieters JLL. 2000. De vinger aan de pols van een gezond Vlaams milieu. Rapportage van een review-commissie van het Project ‘Milieu en Gezondheid’ [in Dutch]. Brussels:Ministerie van de Vlaamse Leefgemeenschap.

Sharpe RM, Skakkebaek NE. 1993. Are oestrogens involved in falling sperm counts and disorders of the male reproductive tract? Lancet 341:1392–1395.

“Sexual Maturation in Relation to Polychlorinated Aromatic Hydrocarbons...”: Den Hond et al.’s Response

The Flemish Government commissioned the Environment and Health Study after a competitive call for research proposals in 1998 (Onderzoek Milieu en Gezondheid 1998). From the outset, this research project included two distinct population-based surveys, one in adolescents (16–18 years of age) and another in adults (21–65 years of age). We assume scientific responsibility only for the project in adolescents.

Our study (which was carried out in 1999) sought to investigate whether, in adolescents, biomarkers can show exposure to and health effects of common environmental pollutants (Staessen et al. 2001). We planned to maximize the gradient in environmental exposure by recruitment of adolescents from a rural control area (Peer, Belgium) and a polluted industrial suburb. For logistic reasons, the adolescents were recruited from and were examined at local schools. Life-long residence in the study areas was a prerequisite for participation. The Flemish Government imposed the selection of the polluted area because of a

long-lasting controversy about the possible impacts on public health of two waste incinerators in Wilrijk and a large nonferrous smelter in Hoboken. Health authorities assumed that our study would settle the local concern about detrimental health outcomes associated with the waste incinerators in Wilrijk. In spite of the imposed selection of the polluted area, our study had the advantage of a clear-cut contrast in exposure between the control and polluted areas, as evidenced by all available records on environmental monitoring.

Molenberghs et al. found the lack of randomness a major shortcoming of our study design. However, they miscalculated the proportion of boys to girls, which was 1.5 (60/40) in Peer, 1.0 (21/21) in Wilrijk, and 2.1 (39/19) in Hoboken. We believe that, in the context of the imposed selection, we maximally exploited the remaining degrees of freedom to avoid serious selection bias. Indeed, participants and nonparticipants had similar age, sex distribution, parental social class, and regional residence. In the industrial suburb they resided at similar distances from the main sources of pollution, which strongly suggests random selection with regard to exposure (Staessen et al. 2001). Although we cannot extrapolate to the general Flemish population, our conclusions are applicable to the adolescents enrolled in our population-based sample.

After our Data Monitoring Committee—which was enlarged by three international experts for peer-review—had discussed the study outcomes on 12 May 2000, our results became available for disclosure to the public and subsequent scientific publication. The health administration then appointed Molenberghs and colleagues as the official external reviewers of our study (Goetghebeur et al. 2000; Cuijpers et al. 2000). These statisticians seem to imply that our conclusions should have been restricted to the feasibility of biomonitoring in adolescents and that we should not have disclosed the unexpected medical findings suggesting relevant environmental health risks. To what extent their criticisms are justified and independent remains open for discussion. In this context, it is interesting to note that, on 25 July 2000, the Flemish Minister of Public Health addressed a letter to the principal investigator of our study (J.A.S.) demanding that disclosure of the findings be postponed for 6 months until the study in adults would be completed. This embargo was lifted on 21 November 2000, a few weeks after the municipal elections.

In contrast to the criticism of Molenberghs et al., we actually based our study on *a priori* hypotheses associating defined biomarkers of exposure with relevant effect markers (Staessen et al. 2001). This is also the case for the specific

article on the endocrine effects of polychlorinated aromatic hydrocarbons (PAHs) (Den Hond et al. 2002) and other recent papers that emanated from the adolescent study, for example, those on the immune (Van Den Heuvel et al. 2002) and genotoxic (Verheyen et al. Unpublished data) effects. Because we were concerned about the reproducibility of the staging of sexual maturation according to Marshall and Tanner's scale (Marshall and Tanner 1969; Marshall and Tanner 1970) and the measurement of testicular volume using Prader's orchidometer (Dörnberger et al. 1987), we mounted a separate validation study of which the results have been reported and discussed elsewhere (Den Hond et al. 2001).

Molenberghs et al. confuse medically assisted conceptions with medically assisted deliveries. Nevertheless, we acknowledge that age is a major determinant of fertility in women. However, the officially published report did not include an age-adjusted analysis of the proportion of medically assisted conceptions (Advoet et al. 1999). The median age in women at first delivery was 27.3 (interquartile range 25.2–29.3) years in Wilrijk and 26.6 (24.2–29.0) years in Flanders. A difference of 0.7 years cannot explain a 65–75% increase in medically assisted conceptions (5.6% in Wilrijk vs. 3.4% in Flanders for singleton births and 59.0% vs. 33.4% for multiple births).

We hope that, with these clarifications, our study is more clearly placed in the perspective of the Flemish Environment and Health Study. We agree with Molenberghs et al. that our findings, although in line with the concept of endocrine disruption and Sharpe and Skakkebaek's hypothesis, require further research. As in all cross-sectional epidemiologic studies, we did not prove causation. However, numerous experimental studies support the causal association between exposure to PCAHs and effects on sexual maturation.

Elly Den Hond

Hypertensie en Cardiovasculaire
Revalidatie Eenheid
Departement Moleculair en
Cardiovasculair Onderzoek
Leuven, Belgium

E-mail: elly.denhond@med.kuleuven.ac.be

Harry Roels

Unité de Toxicologie Industrielle et de
Médecine du Travail
Université catholique de Louvain
Bruxelles, Belgium

Jan A. Staessen

Hypertensie en Cardiovasculaire
Revalidatie Eenheid
Departement Moleculair en
Cardiovasculair Onderzoek
Leuven, Belgium

REFERENCES

- Aelvoet W, Bekaert Y, Dejonghe M. 1999. Gepercipieerd exces aan congenitale afwijkingen en niet-gepercipieerd exces aan meermalen in de buurt van twee huisvuilverbrandingsovens in Neerland. In: *Gezondheidsindicatoren 1997* (Aelvoet W, Fortuin M, Hoofst P, Vanoverloop J, eds). Brussels:Ministerie van de Vlaamse Gemeenschap, 130–141.
- Cuijpers C, Molenberghs G, Passchier W, Pieters J. 2000. De vinger aan de pols van een gezond Vlaams milieu (2) (Dutch). Brussels:Ministerie van de Vlaamse Gemeenschap. Available: http://www.wvc.vlaanderen.be/gezondmilieu/onderzoeken/koepel/rapport_reviewcommissie/pdf/eindrapport.pdf [cited 25 October 2002].
- Den Hond E, Staessen JA, Hoppenbrouwers K, Vanderschueren D, Roels H. 2001. Delayed sexual development in adolescents: author's reply [Letter]. *Lancet* 358:1817.
- Den Hond E, Roels HA, Hoppenbrouwers K, Nawrot T, Thijs L, Vandermeulen C, et al. 2002. Sexual maturation in relation to polychlorinated aromatic hydrocarbons: Sharpe and Skakkebaek's hypothesis revisited. *Environ Health Perspect* 110:771–776.
- Dörnberger V, Dörnberger G. 1987. Vergleichende Volumetrie des menschlichen Hodens unter besonderer Berücksichtigung der Hodensonographie, Praderorchidometer, Schirrenzirkel und Schublehre [Comparative volumetry of human testes using special types of testicular sonography, Praders orchidometer, Schirren's circle and sliding caliper]. *Andrologia* 19:487–496.
- Goetghebeur E, Passchier W, Pieters J. 2000. De vinger aan de pols van een gezond Vlaams milieu. Brussels:Ministerie van de Vlaamse Gemeenschap (Dutch). Available: http://www.wvc.vlaanderen.be/gezondmilieu/onderzoeken/koepel/toxicologie/pdf/addenda_adolescenten/12/pr000607.pdf [cited 25 October 2002].
- Marshall WA, Tanner JM. Variations in the pattern of pubertal changes in girls. *Arch Dis Child* 1969; 44:291–303.
- Marshall WA, Tanner JM. Variations in the pattern of pubertal changes in boys. *Arch Dis Child* 1970; 45:13–23.
- Onderzoek Milieu en Gezondheid. 1998. Ontwikkeling van een gestructureerde aanpak voor vermoedelijke aan het milieu gerelateerde ziekteclusters in Vlaanderen en van een leefmilieu- en epidemiologische monitoring (Dutch). Brussels:Ministerie van de Vlaamse Gemeenschap.
- Staessen JA, Nawrot T, Den Hond E, Thijs L, Fagard R, Hoppenbrouwers K, et al. 2001. Renal function, cytogenetic measurements and sexual development in adolescents in relation to environmental pollutants: a feasibility study of biomarkers. *Lancet* 357:1660–1669.
- Van Den Heuvel RL, Koppen G, Staessen JA, Den Hond E, Verheyen G, Nawrot TS. 2002. Immunologic biomarkers in relation to exposure markers of PCBs and dioxins in Flemish adolescents (Belgium). *Environ Health Perspect* 110:595–600.

Declaring Chemicals “Not Carcinogenic to Humans” Requires Validation, Not Speculation

Regulatory agencies should provide detailed guidelines on how to use mechanistic and epidemiologic data to dismiss positive cancer evidence obtained from studies in experimental animals. Roberts and Ashby (2002) bemoaned that the U.S. Environmental Protection Agency (EPA) does not provide opportunity to conclude that there is no evidence of carcinogenicity from well-conducted and adequately powered epidemiologic studies. In their letter, Roberts and Ashby provided examples for which they claimed that there is enough information to draw valid conclusions of “no effect.” Our examination

of their examples and viewpoints reveals several critical issues that they have ignored, but which must be addressed in any future regulatory agency guidelines that may allow the dismissal of carcinogenic effects in animals based on mechanistic and epidemiologic data.

Roberts and Ashby (2002) appealed to the International Agency for Research on Cancer (IARC) guidelines (IARC 2000a) for evaluating epidemiologic studies to rectify the absence of procedures in the U.S. EPA's guidelines (US EPA 1999) that would allow the conclusion that there is no carcinogenic hazard based on well-conducted epidemiologic studies. However, Roberts and Ashby disregarded the IARC distinction between the classifications "inadequate evidence of carcinogenicity" and "evidence suggesting lack of carcinogenicity" in humans. Requirements for the latter are intentionally stringent, including multiple, mutually consistent, adequately powered studies covering the full range of human exposures that exclude with reasonable certainty bias, confounding, and chance and provide individual and pooled estimates of risk near unity with narrow confidence intervals. In particular, IARC (2000a) cautions that "latent periods substantially shorter than 30 years cannot provide evidence for lack of carcinogenicity."

With the IARC criteria in mind, it is surprising that Roberts and Ashby (2002) pointed to epidemiologic studies on clofibrate as an example, where they claim there is sufficient power to detect a cancer increase and that there are enough data to conclude that there is no carcinogenic effect. Although the World Health Organization trial on prevention of ischemic heart disease with clofibrate includes 208,000 man-years of observation in the mortality update (WHO 1984), the 5-year treatment period plus the 8-year mortality follow-up period is insufficient to adequately determine the extent of drug-related cancer incidences or deaths, or to make a valid conclusion of no risk of cancer. Cancer is a group of diseases that most often have latency periods greater than 20 years for clinical manifestation. Further, because there were only 206 total cancer deaths among the 5,331 men in the clofibrate group (WHO 1984), this study has inadequate power to demonstrate lack of a carcinogenic hazard, which should be evaluated for individual sites. This study also provided no information on cancer risk to women. The meta-analysis by Law et al. (1994) did not provide any estimates of cancer risk for clofibrate, though it was cited by Roberts and Ashby (2002) as part of the negative human evidence on the carcinogenicity of this hypolipidemic drug. Certainly, exposure and follow-up durations, gender differences, and study power are critical issues for regulatory

agencies to address when formulating guidelines for use of epidemiologic studies to support a claim of no evidence of carcinogenicity in humans despite positive cancer data in experimental animals.

On the use of mechanistic data for evaluating chemical carcinogens, we have repeatedly stressed the need for rigorous testing of mechanistic hypotheses (Huff 1995; Melnick et al. 1997). In this regard it is instructive to examine the basis for IARC's conclusions on the relevance of the carcinogenicity of peroxisome proliferators in experimental animals. As noted by Roberts and Ashby (2002), IARC opined that the mechanism of tumor induction by clofibrate [and di(2-ethylhexyl)phthalate (DEHP)] would not be operative in humans. For DEHP, there is sufficient evidence of carcinogenicity in rats and mice, yet IARC downgraded the classification of this chemical from "possibly carcinogenic in humans" to "not classifiable as to its carcinogenicity to humans" (IARC 2000b). Downgrading DEHP was based on the working group's conclusion that in addition to peroxisome proliferation, cell proliferation induced by DEHP is due to activation of the peroxisome proliferator-activated receptor (PPAR α) and that peroxisome proliferation had not been documented in human hepatocyte cultures. However, in considering IARC's evaluation, it is important to recognize that humans possess a functional PPAR α and that the working group noted that effects of DEHP in human liver have not been adequately evaluated. Cell proliferation mediated by PPAR α was considered by the IARC working group to be the critical step in the carcinogenicity of peroxisome proliferators.

Several deficiencies and critical data gaps in the peroxisome proliferation mode-of-action hypothesis have been noted (Melnick 2001). Moreover, recent studies have shown that hepatocyte proliferation and peroxisome proliferation occur by different mechanisms, and that PPAR α -independent effects that occur in Kupffer cells are required for the induction of cell proliferation and suppression of apoptosis (Parzefall et al. 2001; Peters et al. 2001; Rose et al. 1999). Although the demonstration of the important role of Kupffer cells in peroxisome proliferator-mediated effects is fairly recent, it seems odd that this subject was not mentioned in the IARC evaluation of DEHP (IARC 2000b; Melnick 2003). Based on current information, it is clear that peroxisome proliferators cause liver tumors in rodents, but the mechanism of action is not understood (Peters et al. 2001).

Tomatis (2002) warned the occupational and environmental health communities that serious public health consequences might follow if decision-making bodies rely on

untested mechanistic hypotheses that are later shown experimentally to be incorrect. Thus, when formulating guidelines for acceptance of mechanistic hypotheses, we expect regulatory agencies to require rigorous testing and validation before using incomplete data sets to downgrade the categorization of chemical carcinogens. Such guidelines must also address whether differences in mechanistic events among species are truly qualitative rather than quantitative in nature. For quantitative differences, the guidelines should also require information on the range of parameter variability in exposed humans so that sensitive subpopulations are not ignored in these categorizations. The issue is not singularly the adequacy of mechanistic or epidemiologic evidence, but certainty for protecting public health.

Ronald L. Melnick
Freya Kamel
James Huff

National Institute of Environmental
Health Sciences
Research Triangle Park, North Carolina
E-mail: melnickr@niehs.nih.gov

REFERENCES

- Huff J. 1995. Mechanisms, chemical carcinogenesis, and risk assessment: cell proliferation and cancer. *Am J Ind Med* 27:293-300.
- IARC. 2000a. Preamble IARC Monogr Eval Carcinog Risks Hum 77:9-31.
- IARC. 2000b. Di(2-ethylhexyl)phthalate. IARC Monogr Evaluat Carcinog Risks Hum 77: 41-148.
- Law MR, Thompson SG, Wald NJ. 1994. Assessing possible hazards of reducing serum cholesterol. *Br Med J* 308:373-379.
- Melnick RL. 2001. Is peroxisome proliferation an obligatory precursor step in the carcinogenicity of di(2-ethylhexyl)phthalate (DEHP)? *Environ Health Perspect* 109:437-442.
- Melnick RL. 2003. Suppression of crucial information in the IARC evaluation of DEHP [letter]. *Int J Occup Environ Health* 9:84-85.
- Melnick RL, Kohn MC, Huff J. 1997. Weight of evidence versus weight of speculation to evaluate the α 2u-globulin hypothesis. *Environ Health Perspect* 105:904-906.
- Parzefall W, Berger W, Kainzbauer E, Teufelhofer O, Schulte-Hermann R, Thurman RG. 2001. Peroxisome proliferators do not increase DNA synthesis in purified rat hepatocytes. *Carcinogenesis* 22:519-523.
- Peters JM, Rusyn I, Rose ML, Gonzalez FJ, Thurman RG. 2001. Peroxisome proliferator-activated receptor α is restricted to hepatic parenchymal cells, not Kupffer cells: implications for the mechanism of action of peroxisome proliferators in hepatocarcinogenesis. *Carcinogenesis* 21:823-826.
- Roberts R, Ashby J. 2002. Mechanistic and epidemiologic data: when is enough enough? *Environ Health Perspect* 110:A502-A503.
- Rose MI, Rivera CA, Bradford BU, Graves LM, Cattley RC, Schoonhoven R, et al. 1999. Kupffer cell oxidant production is central to the mechanism of peroxisome proliferators. *Carcinogenesis* 20:27-33.
- Tomatis L. 2002. The IARC Monographs program: changing attitudes towards public health. *Int J Occup Environ Health* 8:144-152.
- U.S. EPA. 1999. Guidelines for Carcinogenic Risk Assessment. Draft. Washington, DC: U.S. Environmental Protection Agency.
- WHO. 1984. WHO cooperative trial on primary prevention of ischaemic heart disease with clofibrate to lower serum cholesterol: final mortality follow-up. Report of the Committee of Principal Investigators. *Lancet* 2:600-604.